DOCUMENT RESUME

ED 376 215 TM 022 352

AUTHOR	Webb, Noreen M.
TITLE	Group Collaboration in Assessment: Competing Objectives, Processes, and Outcomes. Project 2.1: Designs for Assessing Individual and Group Problem Solving. Effects of Group Characteristics on Groups and Individual Performance.
INSTITUTION	National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	Nov 93
CONTRACT	R117G10027
NOTE	40p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Cooperation; *Educational Assessment; *Group Dynamics; *Measurement Techniques; *Objectives; *Outcomes of Education; Productivity; Student Evaluation; Test Coaching; Test Construction; Test Use

ABSTRACT

A number of theoretical and practical issues are explored that need to be considered in the design of assessments that use group collaboration to be sure that collaboration works toward, rather than away from, the purpose of the assessment. The traditional purpose of assessment has been to measure individual competence of students in thinking skills and subject matter knowledge. A second purpose can be to measure how well students can perform after they have been given an opportunity to learn from group collaboration. Another purpose is to measure group productivity, and a fourth is to measure student ability to collaborate and function as part of a team. The purpose of the assessment, the goal of group work, and the group processes must be specified clearly and must be measurable. Evaluation procedures and criteria must be consistent with the goal of group work. Students do not automatically know how to work with others, and must be prepared for group assessments. A variety of strategies are suggested for preparation and effective collaboration. (Contains 119 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED376215

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

National Center for Research on
Evaluation, Standards, and Student Testing

Final Deliverable – November 1993

Project 2.1: Designs for Assessing Individual and
Group Problem Solving

Effects of Group Characteristics on Groups
and Individual Performance

Group Collaboration in Assessment: Competing Objectives, Processes, and Outcomes

Noreen Webb, Project Director

U.S. Department of Education
Office of Educational Research and Improvement
Grant No. R117G10027 CFDA Catalog No. 84.117G

Center for the Study of Evaluation
Graduate School of Education
University of California, Los Angeles
Los Angeles, CA 90024-1522
(310) 206-1532

TM022352

BEST COPY AVAILABLE

The work reported herein was supported under the Educational Research and Development Center Program cooperative agreement R117G10027 and CFDA catalog number 84.117G as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the position or policies of the Office of Educational Research and Improvement or the U.S. Department of Education.

GROUP COLLABORATION IN ASSESSMENT: COMPETING OBJECTIVES, PROCESSES, AND OUTCOMES[1]

Noreen M. Webb

Large-scale assessment programs are increasingly starting to include collaborative small-group work instead of, or in addition to, students solving problems or working on tasks individually (e.g., Connecticut's Common Core of Learning Assessment: Baron [1994], Connecticut State Board of Education [1987], Lomask, Baron, Greigh, & Harrison [1992]; California Assessment Program: Awbrey [1992], Bartlett [1992], Pandey [1991]; California Learning Assessment System: Saner, McCaffrey, Stecher, Klein, & Bell [1994]; Oregon State Department of Education: Neuberger [1993]; Shavelson & Baxter [1992]. Recommendations to include small group work have also started appearing in efforts toward developing national standards for assessment (Mathematical Sciences Education Board, National Research Council, 1993).

A major reason for using group work in assessment is to link assessment more closely to the growing emphasis on small-group collaboration and cooperation in classroom instruction (Linn, 1993; Wise & Behuniak, 1993). Recognizing that cooperative and collaborative learning in the classroom can increase student learning and socio-emotional outcomes such as social skills, self-esteem, and attitudes towards others (Bossert, 1988-1989; Slavin, 1990), school districts, state departments of education, and national research organizations have started to mandate the use of cooperative and collaborative learning methods of instruction on a large scale (e.g., California State Department of Education, 1985, 1992; Mathematical Sciences Education Board, National Research Council, 1989). Calls for collaborative small-group work also appear in curriculum standards, for example, to help students develop and communicate ideas and stimulate deeper understanding of the subject matter (National Council of Teachers of Mathematics, 1989).

[1] I wish to thank Jonathan Troper for his help in searching the literature and for his comments on an earlier draft of this report, and Joan Herman for her helpful suggestions.

The most often-cited reasons for using group work in classroom instruction may not coincide with purposes of assessment, however. We need to think carefully about the purpose of an assessment and the effects of collaboration on interpretation of scores from assessments. Most importantly, how group work is used in an assessment should coincide with the purpose of the assessment. This paper discusses a number of theoretical and practical issues that need to be considered in the design of assessments that use group collaboration so that the collaboration will work toward, rather than against, the purpose of the assessment. The subsequent sections of the paper describe (a) four purposes of assessments, (b) how the different purposes represent competing goals of learning from group work versus group productivity, (c) how the group processes that take place in collaborative group work may influence the outcomes of assessment differently depending on the purpose of the assessment (d) why it is important to consider group processes in the design of assessments (e) how the impact of group processes on the outcomes of assessments raises fairness issues for group compositions, and (f) how students can be prepared for assessments to ameliorate equity problems and to optimize student performance.

Purposes of Assessment

Four purposes of assessment are considered here. The traditional purpose of assessment of achievement has been to measure individual competence of students in their thinking skills and subject matter knowledge and expertise. The assumption implicit in most testing programs has been that individual competence is best measured by individuals working alone without assistance from others. From this perspective, any amount of collaboration on the test would invalidate test scores because they would not reflect only the competence of the individual. As will be seen below, an individual's score may be influenced by the expertise of other students in a collaborative group and by the processes that emerge during collaboration, as well as by his or her own competence.

A second purpose of assessment is to measure how well individuals can perform after they are given an opportunity to learn from group collaboration. This alternative perspective on student competence is based on the fact that much learning is constructed in collaboration with others. The social constructivist perspective says that individual competence consists of the knowledge, skills, and understanding that an individual can construct while

working with others. From this perspective, an acceptable measure of individual competence might emerge from individuals collaborating with others to learn how to solve problems that they could not previously solve by themselves (Vygotsky, 1978). Individual competence would be demonstrated not only by the expertise that a student had coming into the test, but also by what the individual was able to learn from the collaborative group experience.

Several large-scale performance assessments build in collaboration that seems to have this purpose. For example, the Connecticut Common Core of Learning Alternative Assessment Program (Baron, 1994) uses collaboration in this way. Students first work individually to provide information about their knowledge and conceptual understanding, then work in collaborative small groups on a complex task or experiment, and finally carry out similar activities individually.

Baron (1994) gives an example assessment from the Connecticut Common Core of Learning Alternative Assessment in Science. In the first part, students individually list the factors that influence yeast's activity in food. In the second part, three-person groups design, carry out, interpret, and summarize (in writing and orally to the class) an experiment investigating the activity of yeast in food. In the third part, students work individually to analyze and critique a report about an experiment that was written by another group. Because the assessment is viewed as an opportunity for students to "deepen their understanding of the concepts and skills being assessed" (Baron, 1994, p. 3), the collaborative part of the assessment was used to give students an opportunity to learn from each other, reflecting the social learning theories of Vygotsky (1978) and others. Students' scores on the individual portions of the test, then, reflect their prior expertise (the first part of the assessment) and a combination of their own competence and what they learned from the group (the third part of the assessment).

A third purpose of assessment is to measure group productivity. What students can accomplish in teams is important to potential employers who are increasingly using work teams to respond to global competition (Hackman, 1990). Assessing students in groups provides information about group productivity and group effectiveness that individual assessment of student skills does not.

The fourth purpose of assessment considered here is to measure students' ability to interact, work, and collaborate with others, and function effectively as a member of a team. Team effectiveness involves many dynamic processes including, for example, coordination, communication, conflict resolution, decision making, problem solving, and negotiation (Salas, Dickinson, Converse, & Tannenbaum, 1992). These and other interpersonal and teamwork skills are now recognized as a central competency that high school students must acquire for workforce readiness (O'Neil, Allred, & Baker, 1992). For example, the Secretary's Commission on Achieving Necessary Skills (SCANS) identified the following interpersonal skills as workforce readiness skills: participating as a member of a team (working cooperatively and productively with others), teaching others new skills, serving clients and customers, exercising leadership, negotiating, and working with diverse groups (O'Neil et al., 1992; U.S. Department of Labor, 1991). Studies conducted by the American Society of Training and Development, the Michigan Employability Skills Task Force, the New York State Education Department, and the National Academy of Sciences have identified similar interpersonal skills (see O'Neil et al., 1992 for a summary).

One way to assess students' interpersonal skills is to observe them collaborating with others and evaluate their ability to work with others and their ability to monitor and shape their own behavior in the collaborative setting (Redding, 1992). Indeed, some achievement assessment programs do measure collaborative behavior such as participating in the group, offering useful ideas, involving others, and communicating clearly, either through self-reports (Connecticut Common Core of Learning Alternative Assessment Program, described in Baron, 1994) or through evaluations by observers (Awbrey, 1992).

Competing Goals of Collaborative Work: Group Productivity vs. Learning

The purposes of assessment described above reveal two different, competing goals of group work: group productivity versus learning from collaborative groupwork. When the goal of group work is group productivity, evaluation focuses on the output of the group. Either the quantity of output, or the quality of the product, or both, may be evaluated (for example, the quantity of ideas for solving a problem, the quality of a solution to a problem). When the goal of

group work is learning on the part of individual members, in contrast, evaluation focuses on learning outcomes, not the quality of the group's performance on the group task.

Not only is what is evaluated different for group productivity and learning contexts, but the group processes leading to good performance may be different in the two contexts. When students know that they will be evaluated on the basis of what they produce, they will (and perhaps should) behave in ways that are different from how they would (and should) behave if their individual learning outcomes are to be evaluated. Behavior that is conducive to producing a high-quality group product may not always be conducive to individual learning, and vice versa.

Because the processes and outcomes of group collaboration may differ depending on whether the goal is individual learning or group productivity, it is important that the purpose of the assessment, the goal of group work, and the group processes supposed to contribute to those goals be specified clearly. The next section describes some of the processes that may take place in group work and discusses whether each process is likely to be beneficial or detrimental for the goals of individual learning and group productivity.

Group Processes and Their Influence on Outcomes of Assessment

A number of different group processes may emerge when students are given a task to perform collaboratively on an assessment. Each group process may be beneficial for one goal, but may undermine other goals. Rarely is a group process desirable for all purposes of assessment simultaneously.

Co-Construction of Ideas

When individuals work collaboratively on a task, they can build on each other's ideas to construct new knowledge, skills, and understanding that they did not have prior to group work (Damon & Phelps, 1989). This process figures prominently in the constructivist perspective of Vygotsky in which interactions with others are internalized and transformed to produce new understanding (Vygotsky, 1978, 1981; Wertsch & Bivens, 1992).

This process will have beneficial effects on performance when the purpose of assessment is either to measure individual learning from collaborative group

work or to measure group productivity. Developing new understanding by building on other students' ideas is a form of learning that students can demonstrate on subsequent individual tests. Building on each other's ideas can also help the group produce a higher quality product or solution to a problem.

Conflict and Controversy

A leading theoretical perspective regarding the mechanisms by which learning occurs in group contexts is Piaget's model of socio-cognitive conflict and learning. Interaction with others may produce discrepancies between a child's views and new information, giving rise to cognitive conflict within the child, and leading the child to try out new ideas. One way that internal cognitive conflict may be manifested in interaction with others is overt conflict or controversy, in which individuals not only recognize that their beliefs are different from those of others but also confront others about their differences. Overt conflict encourages individuals to explain and justify their own positions, raises uncertainties about their beliefs, encourages individuals to seek new information to help resolve their disagreements and arguments, and helps them understand alternative points of view, all of which can promote learning (Brown & Palincsar, 1989; Johnson & Johnson, 1979).

Conversely, conflict avoidance is hypothesized to be detrimental for learning. Suppressing disagreements, whether due to the domination of one group member over others (Brown & Palincsar, 1989) or from social pressures not to challenge others (Johnson & Johnson, 1979), will prevent individuals from recognizing and exploring different perspectives, beliefs, conceptions about the material, and strategies for solving problems.

Conflict and conflict avoidance may operate differently when the goal is to maximize group productivity. Although it may be helpful for a group to spend some time comparing different ideas and strategies for solving a problem or creating a product, spending too much time exploring different options will prevent the group from completing the task. At some point, the group has to select a course of action, even if some members of the group do not agree with the course selected.

Giving and Receiving Elaborated Help

Another process that is beneficial for individual learning but may be detrimental when the goal is group productivity is helping other students understand the material. A number of empirical studies in classroom settings have shown that individuals can learn by explaining material to other group members and, to a lesser extent, by receiving explanations from other students (e.g., Webb, 1989, 1991). Giving explanations, whether to help others or to defend one's ideas, encourages the explainer to reorganize and clarify material, recognize misconceptions and fill in gaps in his or her own understanding, and develop new perspectives and understanding as a result (Bargh & Schul, 1980). Receiving explanations can benefit the receiver by filling in gaps in his or her understanding, correcting misconceptions, and strengthening connections between new information and previous learning (Mayer, 1984; Wittrock, 1990). Students who receive explanations will benefit most if they then use the explanations to solve problems or perform tasks for themselves (Vedder, 1985; Webb, 1992; Webb, Troper, & Fall, 1992).

Failure to seek help, on the other hand, can have negative consequences for individual learning. Students who do not seek help when they need it may never correct their misconceptions or lack of understanding. Students may fail to seek help for many reasons (Nelson-Le Gall, 1981, 1985, 1992; Nelson-Le Gall, Gumerman & Scott-Jones, 1983). Students may fail to realize that they don't understand the material or can't perform the task and need help. Even if students are aware that they need help, they may decide not to seek it for fear of being judged incompetent and undesirable as a work mate, or they may not want to feel indebted to those giving the help or feel obliged to reciprocate the help. Students may believe that help-seeking is undesirable (as a result of classroom norms to be quiet and work alone without disturbing others, or sex-typed role norms that view help-seeking as more appropriate for females than males) or may have received unsatisfactory responses to previous help-seeking attempts (e.g., rebukes, responses that did not aid understanding). Finally, they may believe that no one in the group has the competence or resources to help or that responses will not be helpful, or they may lack motivation or a sense of responsibility to do the work (Cook, 1986; Graham & Barker, 1990; Nelson-Le Gall et al., 1983; Newman, 1990; Webb, 1991).

In contrast to individual learning, the quality of a group's product or solution to a problem does not depend on every group member understanding the material. Although spending some time justifying one's proposed solution may help the group produce a high-quality solution, spending time to ensure that everyone understands how to solve the problem may slow the group down and prevent it from completing the solution. Similarly, seeking help from others may also impede the group's progress.

Equality of Participation

Active involvement by all group members is essential for individual learning from collaborative group work. While it may be possible for some individuals to learn simply by observing others (Bandura, 1986; Schunk, 1987), most theories about learning in group contexts emphasize the active involvement of group members in the construction of new knowledge, skills, and understanding, as discussed above.

Individuals don't always have equal opportunities for involvement and participation in group work. Some members are much more active and influential than others (disproportionate influences of some group members are sometimes called "participation biases," Hoffman, 1965). The most talkative member tends to be extroverted, outgoing, domineering, friendly, energetic, or someone who has a particularly high stake in the outcome of group work (Hoffman, 1965).

In addition to personality characteristics, activity and influence in the group are often related to an individual's status in the group. Status may not necessarily relate to an individual's level of competence. According to expectations states theory, when group members do not have a clear way to judge each others' competence on the task, they will use other "socially evaluated" characteristics (such as ethnic background, race, and gender), called diffuse status characteristics, to form their judgments. These status characteristics will determine relative influence in the group (Berger, Rosenholtz, & Zelditch, 1980). Cohen (1982, 1994; Cohen, Lotan, & Catanzarite, 1990) has found that high-status individuals, especially on ethnic background or racial characteristics, tend to be more active and influential than low-status individuals, while low-status individuals tend to be less assertive and more anxious, talk less, and give fewer suggestions and less information than high-

status individuals. Educational researchers have also found that gender may operate as a status characteristic in small groups, with boys being more active than girls in group interaction, boys' answers prevailing more often than girls' answers, and boys being more successful than girls in obtaining needed help, especially in groups with unequal numbers of girls and boys (Lindow, Wilkinson, & Peterson, 1985; Webb, 1984; Wilkinson, Lindow, & Chiang, 1985).

There is also evidence that the ability composition of the group may produce unequal patterns of participation. Several studies have found that in heterogeneous groups, those in the "middle" may participate less, and have their contributions appreciated less, than they would in more homogeneous groups (Dembo & McAuliffe, 1987; Webb, 1989). One hypothesis for this observed pattern is that a teacher-learner relationship may develop between the "high" and "low" members of heterogeneous groups that excludes those in the middle (Webb, 1991).

When the goal is group productivity, in contrast, it may be most efficient and effective for the most able member or members of the group to solve the problem or complete the task without contributions or participation by less-able group members. Research in organizational and social psychology, for example, shows that the solution to a problem that a group produces may be no better than the solution produced by the best member of that group working alone (Bray, Kerr, & Atkin, 1978; Campbell, 1968; Fox & Lorge, 1962; Hastie, 1986; Hill, 1982; Laughlin & Ellis, 1986; Lorge, Fox, Davitz & Brenner, 1958; Lorge & Solomon, 1955; Marquart, 1955; Street, 1974; for an exception, see Shaw & Ashton, 1976). In fact, on problems with solutions that are difficult to demonstrate or explain, such as estimation of a quantity, the group usually produces a solution that is inferior to the solution that the best member would produce alone and about equal to that of the average of the group (Hastie, 1986; Laughlin & Ellis, 1986). Moreover, the best member working alone typically takes less time to solve the problem than the group does (Hill, 1982). These results suggest that contributions by less-able group members may sometimes interfere with producing a high-quality solution and may slow the group down. Suppressing communication and participation by some members of the group may, then, be an effective strategy for maximizing group productivity.

Social Loafing

Social loafing, or diffusion of responsibility, arises when one or more group members sit back and let others do the work (Karau & Williams, 1993; Slavin, 1990). Individuals may go along for a "free ride" if they believe that their efforts can't or won't be identified or are dispensable (Kerr & Bruun, 1983; Levine & Moreland, 1990). The "free rider" effect can turn into the "sucker effect" when the group members who were doing all of the work discover that they have been taken for a free ride and start to contribute less to group work to avoid being a sucker (Kerr & Bruun, 1983; Salomon & Globerson, 1989). Similarly, Jackson and Harkins (1985) propose that, because people tend to match the efforts of their fellow group members, people will engage in social loafing if they perceive that others are also loafing.

This process is always detrimental for individual learning because a student who is not involved in group work will not benefit from any productive collaboration that does occur. Diffusion of responsibility may not be detrimental for group productivity, however, as was illustrated by a recent study of individual learning and group productivity in mathematics (Webb, 1993). In that study, students solved mathematical problems in heterogeneous small groups and then solved similar problems individually. In small groups, students were encouraged to collaborate and to make sure that everyone in the group understood how to solve the problems. Every student was required to turn in a sheet showing the solutions to the problems. Because groups were heterogeneous in ability (as well as gender and ethnic background), every group had at least one student who could solve the problems and could show others what to do. Consequently, group performance was very high (averaging 95% correct) and all students obtained high scores on work performed in groups.

Not all students actively participated in group work, however. About 60% of the students actively participated, either solving the problems or seeking help from other students to learn how to solve them. The remaining 40% of students copied from other students' papers, asked other students for the procedures to write down without trying to understand them, or did not contribute to group work at all. These latter strategies were not detrimental for group performance, and may even have been beneficial. Had students who did not understand how to solve the problems insisted on having every problem explained to them, the group may not have finished solving the problems.

Although "passive" behavior may have had little consequence for group performance, it had disastrous effects on subsequent performance on the individual test. Students who exhibited "passive" behavior in group work—copying others' papers, being given the procedures to write down, and not contributing—averaged only 19% correct on the individual test. In contrast, students who had actively participated in group work—either solving problems or eliciting help to understand how to solve them—averaged 91% correct on the individual test. Students' behavior in the group was a much more potent predictor of their performance on the individual test than was their ability as measured by a pretest.

In the study just described, no data were collected about why students behaved the way they did. So it is not clear whether students who did not actively participate in group work made a conscious decision not to interfere in the group's work, for fear of slowing down group work or negatively impacting the group's score, or were afraid to ask for help lest others would judge their competence unfavorably, or were simply "social loafers" content to let others do the work. Whatever their reasons for not actively participating, their failure to take responsibility for their learning prevented them from learning from the collaborative groups.

Division of Labor

Division of labor may be beneficial for group productivity, but can be detrimental for individual learning. The most efficient way to complete a task may be to divide the task into parts and assign each group member responsibility for completing one part. The problem with this strategy is that each individual would learn only a portion of the task or material. Recognizing the potential danger inherent in division of labor, educators sometimes modify cooperative learning methods of instruction that use group projects to give every member of the group exposure to all of the material. For example, every group member may be required to study all of the material before selecting a part of it to specialize in for the group product (Maskit, 1986; described in Hertz-Lazarowitz, 1992). Whether this procedure would improve student learning of all of the material or would be feasible or practical on an assessment, however, are unknown.

Implications for the Design of Assessments

The previous discussion suggests that the processes that emerge in group collaboration can strongly influence student performance on assessments. The interpretability of students' scores will depend on how well the processes that emerge in group work correspond to the goals of the assessment. If students perceive that the goal of collaboration is to help them learn from one another, they are more likely to engage in processes beneficial for learning than if they perceive that the goal of collaboration is to produce a quality solution or product, and vice versa. Consequently, it is essential that designers of assessments clarify the purpose of the assessment, the purpose of collaborative work included in the assessment, and the desired group processes. Moreover, these purposes and desired processes must be conveyed clearly to students taking the tests.

Measuring Unassisted Individual Student Competence

If the purpose of assessment is to measure student competence without assistance from others, then nearly all of the group processes just described will invalidate individuals' performance scores. Whenever students participate in collaborative group work, they have the opportunity to gain new knowledge, skills, and understanding and, as a consequence, improve their performance. Saner et al. (1994), for example, examined the effects of working in pairs on subsequent individual performance on a three-part hands-on science assessment (concerning trash disposal and recycling) developed by the California Learning Assessment System. On the first day, students worked individually on short-answer questions about relevant content knowledge. On the second day, they worked in pairs to carry out a scientific investigation. On the third day, they worked individually to interpret the results obtained from their investigation from the previous day and to apply their results. Saner et al. found evidence of carry-over effects from working in pairs, particularly among students who scored low on the first part.

Even a small amount of collaboration may influence students' understanding and performance. Analysis of the Fall 1992 pilot field test of the Connecticut Academic Performance Test developed by the Connecticut State Department of Education (1992) showed that as little as 10 minutes of collaboration had an effect for some students. As part of the 90-minute language arts *Response to Literature* assessment, tenth-grade students read a short story

and were given 10 minutes to share their thoughts about it. On some test forms, students read the story and answered two open-ended questions independently, discussed the story in three-person groups, and then answered five more open-ended questions independently. Analyses of a subset of students' test papers revealed that the 10-minute discussion improved some students' understanding of the story. Some students' responses to questions after the discussion showed new insights about the story, such as understanding special twists in the story or understanding a character's motivation (Wise & Behuniak, 1993). These results show that even a small amount of collaboration would render it impossible to interpret test scores as measuring unassisted student competence.

Measuring Individual Student Learning

If the purpose of the assessment is to measure individual student learning from group collaboration, then the assessment should be designed to encourage processes beneficial for learning, and not focus on group productivity. How can this be done? First, it is important to stress individual accountability for learning in the instructions and structure of the test. The test instructions can focus on individual accountability, and evaluation of student performance can focus on individual, rather than group, portions of the test.

Recognizing the danger that group components of a performance assessment may foster a focus on group productivity and inhibit student accountability for learning, some assessments already highlight individual accountability in the test instructions. The Connecticut Common Core of Learning Assessment Project (Connecticut State Department of Education, 1991), for example, incorporated specific instructions for students that highlighted the importance of learning from group work. Instructions at the beginning of a secondary-level assessment in mathematics told students that they would be held accountable for their own learning: "You will be graded on your individual work, as well as the work of your group." Reminders about the importance of individual learning appeared in the instructions for group work: "Each member of the group will be expected to understand fully the procedures, reasoning, and conclusions reached by the group. These understandings will be assessed in a task to be completed on an individual basis following the groupwork."

To reinforce the importance of individual accountability and to help avoid diffusion of responsibility during group work itself, another common strategy is to require every student to be prepared to summarize and explain their group's work. In the Connecticut assessment described above, for example, instructions told students that "each person should be able to explain fully the conclusions reached by the group" and should be prepared to give an oral presentation on their group's experiment. In the California Assessment Program, a group performance assessment task in world history instructed students to "be ready to present your group's statement orally to the entire class and to answer questions from the evaluators or other students" (Awbrey, 1992).

The kinds of instructions just described do send a message that an important purpose of group work is learning, and may help promote processes that are beneficial for learning. But giving such instructions does not, by itself, eliminate group productivity as a goal. Students may continue to focus on group productivity at the expense of individual learning if the group's project, product, or solution to a problem is to be evaluated, and the evaluation of the group's performance contributes to their scores.

On a test in which both group and individual portions are to be evaluated, the competing goals of learning and group productivity may place students in a quandary. Students may be forced to "choose" between processes that will optimize learning and processes that will optimize group productivity. Seeking clarification and elaboration of ideas and making sure that they understand and agree with the group's solution or strategies for completing the task will help prepare students to answer questions individually and to defend the group's work. But encouraging more competent group members to solve the problem or complete the task without interruption from others or stopping to explain their work may be a better strategy for obtaining a good score on the group product. The only solution to this dilemma may be to completely remove any evaluation of the output of group work.

A second way of promoting group processes beneficial for student learning is to tell students which processes are desirable and why. Informing students that actively participating in group work, asking questions to improve their own understanding, and explaining their thinking, for example, are helpful for learning and will help them obtain higher scores on the test may help steer their behavior in the desired directions.

Measuring Group Productivity

If the purpose of the assessment is to measure group productivity, then that goal should be conveyed clearly to students and the test structured accordingly. An assessment focusing on group productivity would give a group a task to complete, and evaluation would focus on the completed task, not on individual students' contribution to completing the task. Just as a focus on group productivity can interfere with processes conducive to learning from group collaboration, any focus on individual accountability can interfere with processes conducive to maximizing group productivity. For example, a belief that their contributions to the group's product will be evaluated may motivate students to participate even when they believe that the group would operate more efficiently and effectively without their contributions. Although there may be interest in observing the contributions of individual students to describe their participation in group work (see next section), evaluating students on the basis of their individual contributions would be counter to the purpose of measuring group productivity.

Measuring Group Processes

It is also important to clarify the purpose of collaboration in small groups whenever group processes are to be measured and evaluated. Some processes are desirable regardless of the purpose of collaboration, such as staying on the topic, being considerate of other students and their ideas, and expressing one's ideas clearly. But other processes may be desirable for one purpose of collaboration but not the other.

If the purpose of the collaboration is to promote student learning, then one would hope to see students exchanging and discussing everyone's ideas, opinions, and knowledge, helping others, working together rather than separately, actively seeking help and clarification, encouraging others to participate, justifying and explaining one's own ideas, and elaborating on others' ideas.

If the purpose of the collaboration is to use group work to produce a high-quality solution or product, then the processes just described may often be sub-optimal. It may be more effective for students to work separately instead of together, for one or more group members to do most of the work while others contribute little, for one student to take control of group work if group members cannot agree, and to have minimal helping behavior. Penalizing such groups for

failing to work together, failing to involve all students, and failing to help each other may be unreasonable and unfair.

Implications for the Administration and Interpretation of Assessments: Composing "Fair" Groups

The potential impact of group processes on outcomes of assessments also has important implications for test administration. The group processes that emerge during collaborative work depend not only on the goal of group work but also on the composition of the group. Beneficial group processes are more likely to arise in some kinds group compositions than in others, with consequences for the performance of students in these groups. To ensure fairness across groups, then, group composition needs to be considered explicitly in the administration of assessments.

Interestingly, group work is often introduced into performance assessments to increase the fairness of assessments. Neuberger (1993), for example, argued that performance assessments that focus on problem solving and in-depth analysis may put students who have not had an opportunity to learn relevant background information at a disadvantage. So the Oregon State Department of Education incorporated group work into several assessment projects (New Standards Project and experimental activities for Oregon's new Certificate of Initial Mastery) to give students opportunities to learn relevant content that other students may already know (Neuberger, 1993). While allowing students to collaborate with others may give them opportunities to learn background information that they did not have prior to the assessment, it does not automatically ensure fairness. Groups may be "unfair" if they don't give students equal access to intellectual resources of others and if they don't give students equal access to favorable group processes.

This section describes the ways in which some group compositions may be more beneficial than others. Because the composition of the group may have different effects on group processes and outcomes depending on whether the goal is individual or group productivity, the two contexts are discussed separately.

Group Composition and Individual Learning

Past empirical research on learning in collaborative groups shows that the composition of the group in terms of ability or achievement, gender, ethnic background or race, and status influences group processes, the experiences of specific individuals within the group, and students' learning outcomes.

Concerning the mix of ability or achievement level in the group, one might predict that students will learn the most by being in groups with the highest possible level of intellectual resources. Empirical research on learning in collaborative groups in the classroom, however, shows that this straightforward prediction is often false.

Low-achieving students often learn more by being in groups with higher achieving students than in groups with only low-achieving students (e.g., Azmitia, 1988; Bell, Grossen, & Perret-Clermont, 1985; Doise & Mugny, 1984; Hooper & Hannafin, 1988; Tudge & Rogoff, 1989; Webb, 1980), but not always (Ames & Murray, 1981; Glachan & Light, 1982; Hooper, Ward, Hannafin, & Clark, 1989). Studies examining the dynamics of these groups suggest that low achievers will benefit by working with higher achieving students only when they receive the help they need and actively participate in the group's collaboration (Azmitia, 1988; Mugny & Doise, 1978; Webb, 1980).

High-achieving students typically perform well whether they work with other high-achieving students or with lower achieving students (e.g., Azmitia, 1988; Hooper & Hannafin, 1988; Hooper et al., 1989; Skon, Johnson, & Johnson, 1981). One study, however, found that high-achieving students performed best if they did not work with other high achievers. In that study, high achievers learned more in heterogeneous groups than in homogeneous groups because they assumed the role of the teacher and gave more explanations in the former groups than in the latter groups (Webb, 1980).

For medium-achieving students, the effects of group composition are the most complicated. A number of studies have shown that medium achievers in heterogeneous groups with highs and lows may be at a disadvantage because they may be excluded from teacher-learner relationships that develop between high achievers and low achievers and given few opportunities to participate (Dembo & McAuliffe, 1987; see also review by Webb, 1991). Medium achievers participated more actively and learned more when they worked in homogeneous

groups or in groups with a more narrow range of achievement levels (mediums and highs, or mediums and lows; see Webb, 1991). These studies clearly show that being in a group with "more-able" students is not necessarily an advantage for learning.

The gender composition of the group may also influence the group processes that emerge and the subsequent learning of individual group members. A fairly large body of research shows that boys tend to dominate interaction in mixed-gender groups (e.g., Hazelwood et al., 1992; Lockheed & Harris, 1984). One study of predominantly high-achieving white students found that this was true whether girls outnumbered boys or boys outnumbered girls: in both cases, girls directed most of their requests for help to boys, but the boys tended not to give help in return, with detrimental effects on the girls' learning (Webb, 1984). Only in groups with equal numbers of girls and boys did all students participate and learn equally. The gender composition of the group has much less impact on group processes and learning among African-American students, however (Grant, 1986; Webb & Kenderski, 1985), possibly because gender does not influence peer interaction among African-American students as among white students (Grant, 1985).

In general, research on students' status (race, ethnic background, socioeconomic status, gender, or achievement level) shows that low-status students participate less and are less influential than high-status students (Cohen, 1982, 1994; Cohen et al., 1990). Because status is often relative (Cohen, 1994; Cohen et al., 1990), the same individual may have relatively high status in one group and participate actively but may have relatively low status in another group and participate less, with important consequences for learning. Miller and Harrington (1990, 1993) recommend that groups be formed so that status on social characteristics is not salient to group members. For example, they advise against numerically unequal representations of social categories, especially "solo status" such as a single minority student, which may accentuate a group's focus on those social categories, even if it means that some groups are homogeneous. They also recommend combining students who vary on different social characteristics simultaneously, such as a black male, a white male, a black female, and a white female, instead of two black females and two white males.

Group Composition and Group Productivity

Most of the empirical research on the impact of group composition on group performance when the goal of group work is group productivity comes from laboratory studies in social psychological or organizational settings. The typical design has been to compare the performance of groups to the performance of individuals working alone. The results of the studies using intellectual problem-solving tasks (e.g., concept mastery test) generally show that group performance is at least as high as the average of group members' previous individual performance, and is sometimes as high as the most capable group member's previous individual performance (Hastie, 1986; Laughlin & Branch, 1972; Moreland & Levine, 1992).

For example, in one of the most comprehensive studies to systematically compare many different group compositions, Laughlin and Branch (1972) first had 1008 college students take the Terman Concept Mastery Test individually. On the basis of their scores, students were trichotomized as high (H), medium (M), or low (L), and then were assigned to four-person groups to take the test again. Fifteen homogeneous and heterogeneous group compositions were formed: HHHH, HHHM, HHHL, HHMM, HHML, HHLL, HMMM,HMML, HMLL, HLLL, MMMM, MMML, MMLL, MLLL, LLLL.

The results of the Laughlin and Branch study showed that group performance was determined by the level of the highest member of the group and the number of individuals at that ability level. For groups with high-ability members, group performance was proportional to the number of high-ability members in the group: the more high-ability students in the group, the better was the group's performance. The same result appeared in groups in which medium-ability students were highest: group performance was proportional to the number of medium-ability students. Furthermore, the comparisons of individual and group performance showed that individuals benefited from working in groups with persons of higher ability but not in groups with persons of lower ability.

Some social psychologists (e.g., Hastie, 1986; Laughlin & Ellis, 1986) have argued that the nature of the task influences the impact of group composition on group performance. For example, Hastie (1986) and Laughlin and Ellis (1986) proposed that group performance will be determined by the ability of the best

member of the group when the group's task has answers that are obvious and can easily be explained or demonstrated by a competent group member. On tasks with solutions that are nonobvious, Laughlin and Ellis (1986) suggested that having a single competent member may not be sufficient for high group performance because one individual may not be able to convince the group of the correct or best solution. In these situations, it may be necessary to have two competent group members to convince the group of the correct or best response and thus ensure high group performance.

The research just described suggests that, for educational assessments with group productivity as a goal, it will be to a student's advantage to be in a group with the most-capable student or students possible. At a minimum, working in groups with more-capable students will give students an advantage over other students working in groups with less-capable students.

Manipulating Group Composition in Practice

The possibility that some group compositions are more advantageous than others poses tremendous challenges for administrators of assessments. At the very least, to compose groups that are "fair" would require keeping track of the mixtures in each group of ability, gender, race, other social characteristics that may cause status inequalities among students, and previous experience and training in communication skills. Furthermore, previous research has usually studied group composition on one student characteristic at a time. But it is probable that combinations of student characteristics need to be considered simultaneously. For example, determining the optimal group composition for a female student depends on also considering her status on other characteristics, such as ability or achievement and race or ethnic background. Even if we knew which group compositions to recommend as optimal for each student, trying to form groups according to such guidelines would be a logistical nightmare.

Moreover, because the characteristics of students vary from classroom to classroom and from school to school, there is no way to ensure equally fair groups across classrooms and schools. Some classrooms and schools have predominantly low-achieving students. Those students would have no opportunity to work with high-achieving students, unlike students in more heterogeneous classes and schools.

Rather than manipulating group compositions to produce "fair" groups that give students equal access to beneficial group processes that are beneficial for learning, which appears to be an intractable problem, it will be more productive to focus directly on the desired group processes. The next section describes ways in which students can be prepared for collaborative work to increase the incidence of beneficial group processes in all groups, regardless of composition.

Implications for Preparing Students for Group Assessments

As the previous sections of this paper suggest, how well students perform on group assessments will depend in part on the nature of the group processes that emerge. But students do not automatically know how to work effectively with others. Developing basic communications skills for effective collaborative group work takes practice. Even if students have had opportunities to work and learn in groups, it is unlikely that many of them will have developed or practiced such sophisticated skills as giving and receiving elaborated help, engaging in constructive controversy, and building on each others' ideas, for example (Farivar & Webb, 1994a, 1994b). Part of the preparation of students for group assessments, then, should focus on giving students experience in working in groups and building their collaborative skills.

Preparing students to work collaboratively with others will accomplish several things. First, it will help eliminate inequities caused by some students being better prepared for group work than others. Second, it will lessen the impact of group compositions on group processes and performance, and eliminate some of the worry over creating "fair" groups. Training in specific communication skills may help eliminate some of the detrimental processes that would otherwise take place in certain group compositions (such as medium-ability students being ignored in heterogeneous-ability groups). Third, practice with group collaboration and training in specific communication and collaboration skills will help maximize performance in groups. Assessments would then measure what students are capable of accomplishing in groups under favorable conditions.

Students can be prepared for group assessments in many ways. This section describes a variety of approaches ranging from simply providing practice in collaborative group work to structuring group work to require certain kinds of group processes.

Practice in Collaborative Group Work

One way of preparing students for taking group assessments is to give them practice working in groups on activities that have the same goal as future assessments. If students will be expected to learn from collaborative group work, they should have experience in that kind of group setting. If the purpose of the assessments they will take is to measure group productivity, then they should have experience working with others to create a single product or solution to a problem.

Practice will give students experience with collaborative group work so that group assessments will not be totally unfamiliar contexts. Practice will also help students develop communication skills and give them an understanding of what kinds of group processes help them learn from collaborating with others, and what kinds of processes help maximize group productivity.

Promoting Effective Communication Skills

Another way of preparing students for group assessments is to give them training in general interpersonal and teamwork skills that are needed for all kinds of collaborative group work. In recognition that members of a group need to know how to communicate effectively with one another, a number of educational researchers have investigated ways of preparing students to work with others. The resulting programs have focused on developing norms for prosocial behavior and specific helping skills.

To encourage students to listen to others, allow everyone to participate, and resolve disagreements in constructive ways, many cooperative learning methods have students carry out activities to establish norms for cooperative behavior in the classroom and to help students develop and practice communication skills (e.g., Johnson, Johnson, Holubec, & Roy, 1984; Kagan, 1992; Sharan & Sharan, 1976). Some entire programs are built around prosocial development, such as the Child Development Program (Solomon et al., 1985; Solomon, Watson, Schaps, Battistich, & Solomon, 1990). This program strives to promote a cooperative orientation in all classroom activities, and designs experiences to help students develop "autonomy, self-direction, community participation, responsible decision making, being helpful to others, learning to understand and appreciate others, and learning to collaborate with others" (Solomon et al., 1990, p. 236).

Cohen et al.'s (1990) adaptation of *Finding Out / Descubrimiento* (DeAvila & Duncan, 1980) helps develop cooperative norms by preparing teachers in methods of classroom management to increase students' ability to help other people, listen to them, explain and demonstrate how things work, give them what they need, and ask them questions, and by making students responsible for ensuring that everyone gets needed help.

Kagan's (1992) program of cooperative learning also advocates teambuilding, classbuilding, and development of social skills. To help students develop social skills needed for working with others, such as listening, turn taking, helping, praising, polite waiting, encouraging, appreciating, asking for help, staying on task, and resolving conflicts in nonhostile ways, Kagan described roles for students to practice skills, ways of structuring activities to elicit particular social skills, techniques of teacher modeling and reinforcement, and techniques for group reflection and planning.

In the approach of Yackel, Cobb, and Wood (1991; see also Wood & Yackel, 1990), the teacher and students mutually constructed norms in the context of formal group work, including sharing, cooperating, achieving consensus about the answer, justifying one's own work, and understanding other students' procedures.

Equalizing Participation and Influence of Group Members

To prevent high-status students from dominating group discussions or low-status students from being left out of group work, it may be helpful for students to carry out activities designed to equalize participation of all group members. For example, Elizabeth Cohen and her colleagues have developed and tested ways to alter the depressed participation of minority students typically observed in multiracial groups. By manipulating students' expectations of each others' competence, called expectation training, Cohen and colleagues were able to equalize the participation of high-status and low-status students (Cohen, 1973; Cohen & Roper, 1972; Cohen, Lockheed, & Lohman, 1976). In their studies, low-status students received special training on academic and nonacademic tasks and then taught high-status students how to do the tasks, thereby changing high-status students' perceptions of the competence of low-status students. When the treatment consisted only of increasing the competence of low-status students without also manipulating the high-status students' expectations of

low-status students' performance, the usual pattern of high-status dominance in group interaction did not change.

A less expensive and less complicated approach to altering students' expectations of each other's competence is the multi-ability intervention. This approach makes typical status characteristics less salient by raising students' awareness of the multiple skills necessary to do the task (Cohen, 1982). For example, Rosenholtz (1985) developed a six-day curriculum to teach students that multiple abilities, such as visual thinking, intuitive thinking, and reasoning, were needed to solve important problems. By performing group tasks that exemplified each ability, students demonstrated their competence to each other with these alternative abilities. This approach reduced the tendency of high-status students to dominate group activity.

Promoting High-Level Elaboration and Discussion of Ideas

If the purpose of assessment is to measure individual learning from group collaboration, then some of the desirable processes are high-level elaboration and discussion of ideas. Giving students specific training in these processes and structuring the task to control the kinds of interaction that students engage in have successfully promoted desirable processes in empirical classroom research.

Instructions in explaining behavior. One approach to promoting high-level elaboration is to give students specific instruction in how to engage in high-level elaboration. Farivar and Webb (1994a, 1994b) and Swing and Peterson (1982), for example, taught students interpersonal relationship skills and explaining skills to improve their ability to teach other students in small groups. Students were given instruction and practice in giving detailed explanations of how to solve problems instead of giving only answers, and asking explicit, direct, and specific questions, shown by other researchers to be more likely to elicit explanations than vague or general questions (Peterson, Wilkinson, Spinelli, & Swing, 1984; Webb & Kenderski, 1985; Wilkinson, 1985; Wilkinson & Spinelli, 1983).

Role specialization. A popular way of managing and facilitating group work is to assign students to different management roles, each with different prescribed behavior (e.g., the "gatekeeper" to equalize participation in the group, Kagan, 1992). Other roles require students to engage in behavior hypothesized to influence learning directly, such as summarizing and active listening.

Alternately called the "learning leader" (Yager, Johnson, & Johnson, 1985) and the "recaller" (Hythecker, Dansereau, & Rocklin, 1988), the summarizer summarizes the main points of the material. To encourage active processing of material by the non-summarizing students, the "learning listener" or "listener/facilitator" is responsible for detecting errors and omissions in the summary, and must ask questions of the summarizer to help clarify the material. In some cooperative learning methods, the summarizer and listener roles have been incorporated into a complex script for cooperative work (Hythecker et al., 1988; Rocklin, O'Donnell, Dansereau, Lambiotte, Hythecker, & Larson, 1985).

Reciprocal questioning. In reciprocal questioning methods, students ask each other thought-provoking questions about the material to promote elaborated discussion and explanation of ideas (e.g., Fantuzzo, Riggio, Connelly, & Dimeff, 1989; King, 1989, 1990, 1992). Because different students bring different perspectives to bear on the material, the questions generated in a group and the explanations offered in response can expose students to new insights into the material. Answering each others' questions can encourage students to recognize their own misconceptions and gaps in understanding, recognize different viewpoints, seek new information to clarify what puzzles them and to resolve disagreements and differences with others, and reconceptualize and reorganize information to justify their responses or make them clearer to others. And, because they know that other students may evaluate what they say, students may work harder to ask better questions and give more thoughtful answers. These processes, in turn, may increase their understanding and recall of the material.

Explanation prompts. Some researchers have given students specific prompts to encourage them to give elaborated explanations of scientific information and observations in collaborative small groups. Instead of having students ask each other questions to elicit elaboration, these researchers gave students guidelines to use when formulating their own explanations (e.g., comparing one's answer or perspective with someone else's), as well as when responding to others (Coleman, 1992; Palincsar, Anderson, & David, 1993).

Controversy versus concurrence-seeking. Still another approach to shape group interaction is to structure the task to require controversy instead of concurrence-seeking. As described earlier, controversy in a group can increase

learning when students seek additional information to resolve conflicting ideas and information (Johnson & Johnson, 1979).

Conclusions

Group collaboration is an important part of instructional practice for many reasons. By collaborating with other students in the classroom, students can learn from working with each other, they can develop interpersonal skills needed to work effectively with others, they can gain experience working in team settings in which group members share common goals, and they can learn how to work with others to maximize the performance and output of the group.

There may also be an appropriate place for collaborative group work in educational assessment. This paper, however, shows that a number of issues need to be explicitly considered and clarified so that group collaboration will work toward, rather than against, the purpose of the assessment. First, the purpose of the assessment must be made clear: measuring unassisted individual students' competence, measuring how well students can learn by working with others, measuring group productivity, or measuring students' ability to interact and collaborate with others. As this paper has shown, most of these purposes are not compatible.

Second, the goal of group work in the assessment must be clarified. Students, teachers, and administrators should know whether the desired outcome of collaborative work is individual student learning or group productivity. As discussed in this paper, different goals may lead to different group processes and outcomes.

Third, the evaluation procedures and criteria should be made clear. Everyone should understand what will be evaluated: student performance on individual portions of the test, quality of the product arising from group collaboration, or behavior of students as they collaborate with others. And the criteria for "good" performance should be made explicit: high-quality individual student responses, high-quality group products, ability of students to work with others to increase their own and other students' learning, or the ability of students to collaborate with others to achieve a high-quality group product.

Fourth, special care must be taken to make sure that the evaluation procedures and criteria are consistent with the goal of group work and the

purpose of assessment. Individual accountability, whether in terms of performance or group processes, is not consistent with the goal of group productivity. It may not be possible to focus on individual accountability and group productivity simultaneously without sending students mixed messages about how they should behave.

Finally, recommendations for the design and administration of assessments must go hand-in-hand with recommendations for further research. First, future research must systematically observe the effects of the goal of group work (both the stated goal and students' perceptions of the goal) on the group processes that emerge and subsequent student and group performance in assessment settings with different stakes. Second, further research should explore the effects of varying group compositions on processes and outcomes of assessments to determine whether group composition is a source of bias. Third, future research should examine whether preparing students for collaborative group work can help produce desirable group processes in assessment contexts and help reduce the effects of group composition on processes and outcomes.

In conclusion, making sensible and responsible decisions about the design and use of collaborative group work in assessments will require careful consideration of the purpose of the assessment and the goal of group work, and designing task and evaluation criteria that are consistent with the goal of group work and the purpose of the assessment.

References

Ames, G. J., & Murray, F. B. (1981). When two wrongs make a right: Promoting cognitive change by social conflict. Developmental Psychology, 18, 894-897.

Awbrey, M. (1992, September). *History-social science group assessment in California (High school level).* Paper presented at the National Center for Research on Evaluation, Standards, and Student Testing's Conference on "What Works in Performance Assessment," UCLA, Los Angeles, CA.

Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development, 59*, 87-96.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Bargh, J. A., & Schul, Y. (1980). On the cognitive benefit of teaching. Journal of Educational Psychology, 72, 593-604.

Baron, J. B. (1994, April). *Using multi-dimensionality to capture verisimilitude: criterion-referenced performance-based assessments and the ooze factor.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Bartlett, L. D. (1992). Students successfully grapple with lessons of history in innovative group performance tasks. *Social Education, 56*, 101-102.

Bell, N., Grossen, M., & Perret-Clermont, A. N. (1985). Sociocognitive conflict and intellectual growth. In M. W. Berkowitz (Ed.), *Peer conflict and psychological growth* (pp. 41-54). San Francisco: Jossey-Bass.

Berger, J., Rosenholtz, S. J. , & Zelditch, M. (1980). Status organizing processes. *Annual Review of Sociology, 6*, 479-508.

Bossert, S. T. (1989-1990). Cooperative activities in the classroom. *Review of Research in Education, 15*, 225-252.

Bray, R., Kerr, N. L., & Atkin, R. (1978). Effects of group size, problem difficulty, and sex on group performance and member reactions. *Journal of Personality and Social Psychology, 36*, 1224-1240.

Brown, A. L., & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 393-451). Hillsdale, NJ: Erlbaum.

California State Department of Education. (1985). *Mathematics framework for California public schools, kindergarten through grade twelve.* Sacramento, CA: Author.

California State Department of Education. (1992). *Mathematics framework for California public schools, kindergarten through grade twelve.* Sacramento, CA: Author.

Campbell, J. P. (1968). Individual versus group problem solving in an industrial sample. *Journal of Applied Psychology, 52,* 205-210.

Cohen, E. G. (1973). Modifying the effects of social structure. *American Behavioral Scientist, 16,* 861-879.

Cohen, E. G. (1982). Expectation states and interracial interaction in school settings. *American Review of Sociology, 8,* 209-235.

Cohen, E. G. (1994) Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64,* 1-36.

Cohen, E. G., Lockheed, M. E., & Lohman, M. R. (1976). The center for interracial cooperation: A field experiment. Sociology of Education, 59, 47-58.

Cohen, E. G., Lotan, R., & Catanzarite, L. (1990). Treating status problems in the cooperative classroom. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 203-230). New York: Praeger.

Cohen, E. G., & Roper, S. (1972). Modification of interracial interaction disability: An application of status characteristics theory. *American Sociological Review, 37,* 643-657.

Coleman, E. B. (1992). *Facilitating conceptual understanding in science: A collaborative explanation-based approach.* Unpublished doctoral dissertation, University of Toronto, Toronto, Canada.

Connecticut State Board of Education. (1987). *Common Core of Learning.* Hartford: Author.

Connecticut State Board of Education. (1991). *Common Core of Learning Assessment [Math / The Federal Debt].* Hartford: Author.

Connecticut State Board of Education. (1992). *Connecticut Academic Performance Test.* Hartford: Author.

Cook, S. W. (1986, April). *Reactions to helping and being helped in interracial cooperative groups: Effects on respect and liking for group members.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Damon, W., & Phelps, E. (1989). Critical distinctions among three methods of peer education. *International Journal of Educational Research, 13,* 9-19.

DeAvila, E. A., & Duncan, S. E. (1980). *Finding out/Descubrimiento*. Corte Madera, CA: Linguametrics Group.

Dembo, M. H., & McAuliffe, T. J. (1987). Effects of perceived ability and grade status on social interaction and influence in cooperative groups. *Journal of Educational Psychology, 79*, 415-423.

Doise, W., & Mugny, G. (1984). *The social development of the intellect*. Oxford: Pergamon Press.

Fantuzzo, J. W., Riggio, R. E., Connelly, S., & Dimeff, L. A. (1989). Effects of reciprocal peer tutoring on academic achievement and psychological adjustment: A component analysis. *Journal of Educational Psychology, 81*, 173-177.

Farivar, S. H., & Webb, N. M. (1994a). Are your students prepared for group work? *Middle School Journal, 25*, 51-54.

Farivar, S. H., & Webb, N. M. (1994b).Helping and getting help—essential skills for effective group problem solving. *Arithmetic Teacher, 41*, 521-525.

Fox, D. J., & Lorge, I. (1962). The relative quality of decisions written by individuals and by groups as the available time for problem solving is increased. *Journal of Social Psychology, 57*, 227-242.

Glachan, M., & Light, P. H. (1982). Peer interaction and learning: Can two wrongs make a right? In G. Butterworth & P. Light (Eds.), *Social cognition: Studies of the development of understanding* (pp. 238-262). Chicago: University of Chicago Press.

Graham, S., & Barker, G. (1990). The down side of help: An attributional-developmental analysis of helping behavior as a low-ability cue. *Journal of Educational Psychology, 82*, 7-14.

Grant, L. (1985). Race-gender status, classroom interaction, and children's socialization in elementary school. In L. C. Wilkinson & C. B. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 57-78). Orlando, FL: Academic Press.

Grant, L. (1986, April). *Classroom peer relationships of minority and nonminority students*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Hackman, J. R. (1990). *Groups that work (and those that don't): Creating conditions for effective teamwork*. San Francisco: Jossey-Bass.

Hastie, R. (1986). Experimental evidence on group accuracy. In B. Grofman & G. Guillermo (Eds.), *Information pooling and group decision making*. Greenwich, CT: JAI Press.

Hazelwood, C. C., Roth, K. J., Hasbach, C., Hoekwater, E., Ligett, C., Lindquist, B., Peasley, K., & Rosaen, C. (1992). *Gender and discourse: The unfolding "living text" of a science lesson* (Tech. Rep. No. 60). East Lansing: Michigan State University, Center for the Learning and Teaching of Elementary Subjects.

Hertz-Lazarowitz, R. (1992). Understanding interactive behaviors: Looking at six mirrors of the classroom. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: Theoretical anatomy of group learning* (pp. 71-101). New York: Cambridge University Press.

Hill, G. W. (1982). Group versus individual performance: Are n + 1 heads better than one? *Psychological Bulletin, 91*(3), 517- 539.

Hoffman, L. R. (1965). Group problem solving. In L. Berkowitz (Ed.), *Advances in experimental social psychology (Vol. 2)*. New York: Academic Press.

Hooper, S., & Hannafin, M. J. (1988). Cooperative CBI: The effects of heterogeneous versus homogeneous grouping on the learning of progressively complex concepts. *Journal of Educational Computing Research*, *4*, 413-424.

Hooper, S., Ward, T. J., Hannafin, M. J., & Clark, H. T. (1989). The effects of aptitude composition on achievement during small group learning. *Journal of Computer-Based Instruction*, *16*, 102-109.

Hythecker, V. I., Dansereau, D. F., & Rocklin, T. R. (1988). An analysis of the processes influencing the structured dyadic learning environment. *Educational Psychologist*, *23*, 23-27.

Jackson, J. M., & Harkins, S. G. (1985). Equity in effort: An explanation of the social loafing effect. *Journal of Personality and Social Psychology*, *49*, 1199-1206.

Johnson, D. W., & Johnson, R. T. (1979). Conflict in the classroom: Controversy and learning. *Review of Educational Research*, *49*, 51-70.

Johnson, D. W., Johnson, R. T., Holubec, E. J., & Roy, P. (1984). *Circles of learning*. Alexandria, VA: Association for Supervision and Curriculum Development.

Kagan, S. (1992). *Cooperative learning*. San Juan Capistrano, CA: Resources for Teachers (27128 Paseo Espada, Suite 622).

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, *65*, 681-706.

Kerr, N. L., & Bruun, S. E. (1983). Dispensability of member effort and group motivation losses: Free rider effects. *Journal of Personality and Social Psychology, 44*, 78-94.

King, A. (1989). Effects of self-questioning training on college students' comprehension of lectures. *Contemporary Educational Psychology, 14*, 366-381.

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal, 27*, 664-687.

King, A. (1992). Facilitating elaborative learning through guided student-generated questioning. *Educational Psychologist, 27*, 111-126.

Laughlin, P. R., & Branch, L. G. (1972). Individual versus tetradic performance on a complementary task as a function of initial ability level. *Organizational Behavior and Human Performance, 8*, 201-216.

Laughlin, P. R., & Ellis, A. L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology, 22*, 177-189.

Levine, J. M., & Moreland, R. L. (1990). Progress in small group research. *Annual Review of Psychology, 41*, 585-634.

Lindow, J. A., Wilkinson, L. C., & Peterson, P. L. (1985). Antecedents and consequences of verbal disagreements during small-group learning. *Journal of Educational Psychology, 77*(6), 658-667.

Linn, R. L. (1993). Educational assessment: Expanded expectations and challenges. *Educational Evaluation and Policy Analysis, 15*, 1-16.

Lockheed, M. E., & Harris, A. M. (1984). Cross-sex collaborative learning in elementary classrooms. *American Educational Research Journal, 21*, 275-294.

Lomask, M., Baron, J., Greigh, J., & Harrison, C. (1992, March). *ConnMap: Connecticut's use of concept mapping to assess the structure of students' knowledge of science*. A symposium presented at the annual meeting of the National Association of Research in Science Teaching, Cambridge, MA.

Lorge, I., Fox, D., Davitz, J., & Brenner, M. (1958). A survey of studies contrasting the quality of group performance and individual performance, 1920-1957. *Psychological Bulletin, 55*, 337-371.

Lorge, I., & Solomon, H. (1955). Two models of group behavior in the solution of eureka-type problems. *Psychometrika, 20*(2), 139-148.

Marquart, D. I. (1955). Group problem solving. *Journal of Social Psychology, 41*, 102-113.

Maskit, D. (1986). *Cooperative teaching and learning in adult learning.* Unpublished master's thesis, University of Haifa, Israel, School of Education. (In Hebrew).

Mathematical Sciences Education Board, National Research Council. (1989). *Everybody counts: A report to the nation of the future of mathematics education*. Washington, DC: National Academy Press.

Mathematical Sciences Education Board, National Research Council. (1993). *Measuring Up: Prototypes for mathematics assessment*. Washington, DC: National Academy Press.

Mayer, R. E. (1984). Aids to prose comprehension. *Educational Psychologist, 19*, 30-42.

Miller, N., & Harrington, H. J. (1990). A situational identity perspective on cultural diversity and teamwork in the classroom. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 39-76). New York: Praeger.

Miller, N., & Harrington, H. J. (1993). Social categorization and intergroup acceptance: Principles for the design and development of cooperative learning teams. In R. Hertz-Lazarowitz and N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 203-227). New York: Cambridge University Press.

Moreland, R. L., & Levine, J. M. (1992). The composition of small groups. *Advances in Group Processes, 9*, 237-280.

Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology, 8*, 181-192.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

Nelson-Le Gall, S. (1981). Help-seeking: An understudied problem-solving skill in children. *Developmental Review, 1*, 224-246.

Nelson-Le Gall, S. (1985). Help-seeking behavior in learning. In E. V. Gordon (Ed.), *Review of research in education* (Vol. 12, pp. 55-90). Washington, DC: American Educational Research Association.

Nelson-Le Gall, S. (1992). Children's instrumental help-seeking: Its role in social acquisition and construction of knowledge. In R. Hertz-Lazarowitz and N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 49-70). New York: Cambridge University Press.

Nelson-Le Gall, S., Gumerman, R. A., & Scott-Jones, D. (1983). Instrumental help-seeking and everyday problem-solving: A developmental perspective. *New directions in helping* (Vol. 2, pp. 265-283). New York: Academic Press.

Neuberger, W. (1993, September) . *Making group assessments fair measures of students' abilities*. Paper presented at the National Center for Research on Evaluation, Standards, and Student Testing's Conference on "Assessment Questions: Equity Answers," UCLA, Los Angeles, CA.

Newman, R. S. (1990). Children's help-seeking in the classroom: The role of motivational factors and attitudes. *Journal of Educational Psychology, 82*, 71-80.

O'Neil, H. F., Allred, K., & Baker, E. L. (1992). *Measurement of workforce readiness: Review of theoretical frameworks* (CSE Tech. Rep. No. 343). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Palincsar, A. S., Anderson, C., & David, Y. M. (1993). Pursuing scientific literacy in the middle grades through collaborative problem solving. *The Elementary School Journal, 93* (5), 643-658.

Pandey, T. (1991). *A sampler of mathematics assessment*. Sacramento: California Department of Education.

Peterson, P. L., Wilkinson, L. C., Spinelli, F., & Swing, S. R. (1984). Merging the process-product and the sociolinguistic paradigms: Research on small-group process. In P. L. Peterson, L. C. Wilkinson, & M. Hallinan (Ed.), *The social context of instruction* (pp. 126-152). Orlando, FL: Academic Press.

Redding, N. (1992). Assessing the big outcomes. *Educational Leadership, 49*, 49-53.

Rocklin, T., O'Donnell, A., Dansereau, D. F., Lambiotte, J. G., Hythecker, V., & Larson, C. (1985). Training learning strategies with computer-aided cooperative learning. *Computers and Education, 9*, 67-71.

Rosenholtz, S. J. (1985). Modifying status expectations in the traditional classroom. In J. Berger, & M. Zelditch (Eds.), *Status, rewards, and influence* (pp. 445-470). San Francisco, CA: Jossey-Bass.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 132-160). Norwood, NJ: Ablex.

Salomon, G., & Globerson, T. (1989). When teams do not function the way they ought to. *International Journal of Educational Research, 13*, 89-99.

Saner, H., McCaffrey, D., Stecher, B., Klein, S., & Bell, R. (1994) *The effects of working in pairs in science performance assessments.* Manuscript submitted for publication. Santa Monica, CA: RAND.

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research, 57,* 149-174.

Sharan, S., & Sharan, Y. (1976). *Small-group teaching.* Englewood Cliffs, NJ: Educational Technology Publications.

Shavelson, R. J., & Baxter, G. P. (1992). What we've learned about assessing hands-on science. *Educational Leadership, 49,* 20-25.

Shaw, M. E., & Ashton, N. (1976). Do assembly effects occur on disjunctive tasks? *Bulletin of the Psychonomic Society, 8,* 469- 471.

Skon, L., Johnson, D. W., & Johnson, R. T. (1981). Cooperative peer interaction versus individual competition and individualistic efforts: Effects on the acquisition of cognitive reasoning strategies. *Journal of Educational Psychology,* 73, 83-92.

Slavin, R. E. (1990). *Cooperative learning: Theory, research, and practice.* Englewood Cliffs, NJ: Prentice-Hall

Solomon, D., Watson, M., Battistich, V., Schaps, E., Tuck, P., Solomon, J., Cooper, C., & Ritchey, W. (1985). A program to promote interpersonal consideration and cooperation in children. In R. Slavin, S. Sharan, S. Kagan, R. Hertz-Lazarowiz, C. Webb, & R. Schmuck (Eds.), *Learning to cooperate, cooperating to learn* (pp. 371-402). New York: Plenum.

Solomon, D., Watson, M., Schaps, E., Battistich, V., & Solomon, J. (1990). Cooperative learning as part of a comprehensive classroom program designed to promote prosocial development. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 231-260). New York: Praeger.

Street, W. R. (1974). Brainstorm by individuals, coacting and interacting groups. *Journal of Applied Psychology, 59*(4), 433- 436.

Swing, S. R., & Peterson, P. L. (1982). The relationship of student ability and small-group interaction to student achievement. American Educational Research Journal, 19, 259-274.

Tudge, J., & Rogoff, B. (1989). Peer influences on cognitive development: Piagetian and Vygotskian perspectives. In M. H. Bornstein & J. S. Bruner (Eds.), *Interaction in human development* (pp. 17-40). Hillsdale, NJ: Lawrence Erlbaum Associates.

U.S. Department of Labor. (1991, June). *What work requires of schools. A SCANS report for America 2000.* Washington, DC: U.S. Department of Labor, The Secretary's Commission on Achieving Necessary Skills (SCANS).

Vedder, P. (1985). *Cooperative learning. A study on processes and effects of cooperation between primary school children.* Westerhaven Groningen, Netherlands: Rijkuniversiteit Groningen.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds. and trans.). Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1981). The genesis of higher mental functioning. In J. V. Wertsch (Ed.), *The concept of activity in Soviet psychology* (pp. 144-188). Armonk, NY: Sharpe.

Webb, N. M. (1980). A process-outcome analysis of learning in group and individual settings. *Educational Psychologist, 15*, 69-83.

Webb, N. M. (1984). Sex differences in interaction and achievement in cooperative small groups. *Journal of Educational Psychology, 76*, 33-34.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13*, 21-40.

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education, 22*, 366-389.

Webb, N. M. (1992). Testing a theoretical model of student interaction and learning in small groups. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 102-119). New York: Cambridge University Press.

Webb, N. M. (1993).Collaborative group versus individual assessment in mathematics: Processes and outcomes. *Educational Assessment, 1*, 131-152.

Webb, N. M., & Kenderski, C. M. (1985). Gender differences in small group interaction and achievement in high-achieving and low-achieving classrooms. In L. C. Wilkinson & C. B. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 209-226). Orlando, FL: Academic Press.

Webb, N. M., Troper, J., & Fall, J. R. (1992, April). *Effective helping behavior in cooperative small groups.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Wertsch, J. V., & Bivens, J. (1992). The social origins of individual mental functioning: Alternatives and perspectives. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 14* (2), 35-44.

Wilkinson, L. C. (1985). Communication in all-student mathematics groups. *Theory Into Practice, 24*(1), 8-13.

Wilkinson, L. C., Lindow, J., & Chiang, C.P. (1985). Sex differences and sex segregation in students' small-group communication. In L. C. Wilkinson & C. B. Marrett (Eds.). *Gender influences in classroom interaction* (pp. 185-208). Orlando, FL: Academic Press.

Wilkinson, L. C., & Spinelli, F. (1983). Using requests effectively in peer-directed instructional groups. *American Educational Research Journal, 20*, 479-502.

Wise, N., & Behuniak, P. (1993, April). *Collaboration in student assessment.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Wittrock, M. C. (1990). Generative processes of comprehension. *Educational Psychologist, 24*, 345-376.

Wood, T., & Yackel, E. (1990). The development of collaborative dialogue within small group interactions. In L. P. Steffe & T. Wood (Eds.), *Transforming early childhood mathematics education: An international perspective* (pp. 244-252). Hillsdale, NJ: Erlbaum.

Yackel, E., Cobb, P., & Wood, T. (1991). Small-group interactions as a source of learning opportunities in second-grade mathematics. *Journal for Research in Mathematics Education, 22*, 390-408.

Yager, S., Johnson, D. W., & Johnson, R. T. (1985). Oral discussion, group-to-individual transfer, and achievement in cooperative learning groups. *Journal of Educational Psychology, 77*, 60-66.